
Phase Transitions in Phase Retrieval

Dustin G. Mixon

Department of Mathematics and Statistics
Air Force Institute of Technology
Wright-Patterson AFB, Ohio 45433

Summary. Consider a scenario in which an unknown signal is transformed by a known linear operator, and then the pointwise absolute value of the unknown output function is reported. This scenario appears in several applications, and the goal is to recover the unknown signal – this is called phase retrieval. Phase retrieval has been a popular subject of research in the last few years, both in determining whether complete information is available with a given linear operator, and in finding efficient and stable phase retrieval algorithms in the cases where complete information is available. Interestingly, there are a few ways to measure information completeness, and each way appears to be governed by a phase transition of sorts. This chapter will survey the state of the art with some of these phase transitions, and identify a few open problems for further research.

Key words: phase retrieval, phase transition, informationally complete, full spark, almost injectivity, unit norm tight frames

1 Introduction

Various applications feature an inverse problem called *phase retrieval*, in which one is given the pointwise absolute value of a known linear transformation of the desired signal. Note that such information will never completely determine the unknown signal since, for example, negating the input will lead to the same output. Indeed, the best one can do is recover $\{\omega x : |\omega| = 1\}$ if x is the unknown signal, but this global phase factor of ambiguity tends not to be an issue in application (for example, in some applications, the true signal is actually nonnegative everywhere, thereby removing all ambiguity). What follows is a brief overview of some of the applications of phase retrieval:

- *Coherent diffractive imaging.* This is a technique to image a nanoscale object by striking it with a highly coherent beam of X-rays to produce a diffraction pattern. The diffraction pattern is the Fourier transform of the object, but only the intensity of the pattern can be physically measured (by counting photons in different regions) [11, 33, 39, 41].

- *Optics.* This application enjoys various instances of phase retrieval: (1) When imaging a star by a lens, one receives the pointwise absolute value of the Fourier transform of the desired pupil distribution [53]. (2) For a high-resolution image, one can apply interferometric techniques to approximate the spatial coherence function (which is the Fourier transform of the desired object intensity), though the phase of this function is difficult to estimate accurately, so it is discarded [21]. (3) Soon after NASA launched its Hubble Space Telescope, they discovered that the primary mirror in the telescope suffered from a large spherical aberration; the extent of this aberration was established by determining the pupil function from the intensity of its Fourier transform (the point spread function) [29].
- *Quantum state tomography.* When measuring a pure state (i.e., a unit vector) x with a positive operator-valued measure of rank-1 elements $\{\varphi_n \varphi_n^*\}_{n=1}^N$ (i.e., the outer products of Parseval frame elements), the random outcome Z of the measurement has a distribution given by $\Pr(Z = n) = |\langle x, \varphi_n \rangle|^2$. As such, repeated measurements will produce an empirical estimate of the distribution, which is the pointwise absolute value (squared) of a linear transformation of the desired signal x [30, 31, 35].
- *Speech processing.* One common method of speech signal denoising is to take the short-time Fourier transform (STFT) and perform a smoothing operation on the magnitudes of the coefficients [5]. Instead of inverting the STFT with the noisy phases, one can recover the denoised version by phase retrieval [49].

Though there are many applications of phase retrieval, the task is often impossible; in particular, discarding the phases of the Fourier transform is not at all injective. This fact has led many researchers to invoke a priori knowledge of the desired signal, since injectivity might be gotten by restricting to a smaller signal class. For example, for the optics applications, the pupil distribution is only supported within the aperture of the optical system, and so this compact-support constraint combined with the intensity measurements might uniquely determine the desired signal. Introducing such information has led to various ad hoc phase retrieval algorithms, and while they have found some success (e.g., in correcting the Hubble Space Telescope), such algorithms often fail to work unexpectedly. Overall, this route has yet to produce algorithms with practical performance guarantees.

Thankfully, an alternative route was introduced in 2006 by Balan, Casazza and Edidin [5]: Seek injectivity, not by restricting to a smaller signal class, but rather by designing a larger ensemble of measurement vectors. Unbeknownst to Balan et al. at the time, this idea had already been in the air in the quantum community (for quantum state tomography [30, 31]), but posing the idea to the signal processing community led to a flurry of research in search of practical phase retrieval guarantees [1, 3, 4, 7, 13, 14, 15, 16, 23, 26, 51, 52]. One popular method called *PhaseLift* recasts phase retrieval as a semidefinite program [13, 14, 15, 16], another called *PhaseCut* reformulates it in terms of

MaxCut [51, 52], and yet another uses the polarization identity along with angular synchronization to quickly solve certain instances [1, 7]. In this same line of research, a new methodology for coherent diffractive imaging emerged [13]: Instead of taking a single exposure and attempting phase retrieval with possibly incomplete information, take multiple exposures of the same object with different masks or diffraction gratings. Not only can such a process produce complete information, there are also provably efficient (and apparently stable) phase retrieval algorithms for this setting [7, 15]. Considering phase retrieval has a wide range of applications, it would be interesting to find other areas to apply this philosophy of taking more measurements to obtain injectivity.

Another effect of the paper [5] by Balan, Casazza and Edidin was the community's desire for a deeper understanding of injectivity for phase retrieval. In particular, what are the conditions for injectivity, and how many measurements are required? Understanding this will help in determining how many (and what type of) exposures are necessary for coherent diffractive imaging; also, such phase retrieval results can be directly interpreted as fundamental limits of quantum state tomography. The purpose of this chapter is to survey several results along these lines, and to identify some remaining open problems. Section 2 focuses on injectivity in both the real and complex cases, and then the third section considers a relaxed version of injectivity called *almost injectivity*. First, we discuss the notation we use throughout this chapter as well as some preliminaries:

1.1 Notation and Preliminaries

Given a collection of vectors $\Phi = \{\varphi_n\}_{n=1}^N$ in $V = \mathbb{R}^M$ or \mathbb{C}^M , we will identify such a collection with the $M \times N$ matrix whose columns form the collection. Consider the intensity measurement process defined by

$$(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2.$$

Note that $\mathcal{A}(x) = \mathcal{A}(y)$ whenever $y = cx$ for some scalar c of unit modulus. As such, the mapping $\mathcal{A}: V \rightarrow \mathbb{R}^N$ is necessarily not injective. To resolve this (technical) issue, throughout this chapter, we consider sets of the form V/S , where V is a vector space and S is a multiplicative subgroup of the field of scalars. By this notation, we mean to identify vectors $x, y \in V$ for which there exists a scalar $c \in S$ such that $y = cx$; we write $y \equiv x \bmod S$ to convey this identification. Most (but not all) of the time, V/S is either $\mathbb{R}^M/\{\pm 1\}$ or \mathbb{C}^M/\mathbb{T} (here, \mathbb{T} is the complex unit circle), and we view the intensity measurement process as a mapping $\mathcal{A}: V/S \rightarrow \mathbb{R}^N$; it is in this way that we will consider the measurement process to be injective.

As the title suggests, the focus of this chapter is phase transitions in phase retrieval. As an example of a phase transition, consider what it takes for a collection of members of a vector space V to span V . Certainly, no collection of size less than the dimension of V has a chance of spanning, and we expect

most collections of size at least the dimension to span. As such, we might say that the notion of spanning V exhibits a phase transition at the dimension of V . We make this definition explicit in the following:

Definition 1. Let $A[\Phi; \mathbb{F}^{M \times N}]$ be a statement about a matrix $\Phi \in \mathbb{F}^{M \times N}$, and consider a function $f: \mathbb{N} \rightarrow \mathbb{N}$. We say $A[\Phi; \mathbb{F}^{M \times N}]$ exhibits a phase transition at $N = f(M)$ if for each $M \geq 2$,

- (a) $A[\Phi; \mathbb{F}^{M \times N}]$ does not hold whenever $N < f(M)$, and
- (b) for each $N \geq f(M)$, there exists an open, dense subset $S \subseteq \mathbb{F}^{M \times N}$ such that $A[\Phi; \mathbb{F}^{M \times N}]$ holds for every $\Phi \in S$.

Based on experience, both parts (a) and (b) of a phase transition are established by first studying necessary and sufficient conditions for the property of interest $A[\Phi; \mathbb{F}^{M \times N}]$. For part (b) in particular, algebraic geometry consistently plays a key role. Viewing $\Phi \in \mathbb{F}^{M \times N}$ as a point in real Euclidean space, consider the collection of Φ 's for which $A[\Phi; \mathbb{F}^{M \times N}]$ does not hold. If there exists a real, nontrivial algebraic variety that contains these points (which is often the case), then every point Φ in the complement of the variety (which is open and dense in $\mathbb{F}^{M \times N}$) satisfies $A[\Phi; \mathbb{F}^{M \times N}]$. As such, for part (b), it suffices to identify the appropriate variety.

Throughout this chapter, we will continually follow a certain procedure for studying phase transitions. Given a property $A[\Phi; \mathbb{F}^{M \times N}]$, we start by studying various necessary and sufficient conditions for that property. We then attempt to prove a phase transition $N = f(M)$ using the conditions available. Later, we consider explicit constructions of $M \times f(M)$ matrices which satisfy $A[\Phi; \mathbb{F}^{M \times f(M)}]$; these minimal constructions are certainly mathematically interesting, and they are also optimal measurement designs for applications like quantum state tomography.

2 Injectivity

In this section, we study the phase transition for injectivity. As we will see, this phase transition is much better understood in the real case than in the complex case, and the distinction is rather interesting. It is highly recommended that the reader enjoys the real case before venturing into the complex case.

2.1 Injectivity in the Real Case

We start by defining the important concepts of this subsection:

Definition 2.

- (a) $\text{Inj}[\Phi; \mathbb{R}^{M \times N}]$ denotes the statement that $\mathcal{A}: \mathbb{R}^M / \{\pm 1\} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$ is injective, where $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$.

(b) $\text{CP}[\Phi; \mathbb{F}^{M \times N}]$ denotes the statement that $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{F}^M$ satisfies the complement property: for every $S \subseteq \{1, \dots, N\}$, either $\{\varphi_n\}_{n \in S}$ or $\{\varphi_n\}_{n \in S^c}$ spans \mathbb{F}^M .

Interestingly, the complement property characterizes injectivity in the real case. Much of the work in this chapter was inspired by the proof of the following result:

Theorem 1 (Theorem 2.8 in [5]). $\text{Inj}[\Phi; \mathbb{R}^{M \times N}]$ if and only if $\text{CP}[\Phi; \mathbb{R}^{M \times N}]$. In words, \mathcal{A} is injective if and only if Φ satisfies the complement property.

Proof. We will prove both directions by obtaining the contrapositives.

(\Rightarrow) Assume Φ does not satisfy the complement property. Then there exists $S \subseteq \{1, \dots, N\}$ such that neither $\{\varphi_n\}_{n \in S}$ nor $\{\varphi_n\}_{n \in S^c}$ spans \mathbb{R}^M . This implies that there are nonzero vectors $u, v \in \mathbb{R}^M$ such that $\langle u, \varphi_n \rangle = 0$ for all $n \in S$ and $\langle v, \varphi_n \rangle = 0$ for all $n \in S^c$. For each n , we then have

$$\begin{aligned} |\langle u \pm v, \varphi_n \rangle|^2 &= |\langle u, \varphi_n \rangle|^2 \pm 2 \text{Re} \langle u, \varphi_n \rangle \overline{\langle v, \varphi_n \rangle} + |\langle v, \varphi_n \rangle|^2 \\ &= |\langle u, \varphi_n \rangle|^2 + |\langle v, \varphi_n \rangle|^2. \end{aligned}$$

Since $|\langle u + v, \varphi_n \rangle|^2 = |\langle u - v, \varphi_n \rangle|^2$ for every n , we have $\mathcal{A}(u + v) = \mathcal{A}(u - v)$. Moreover, u and v are nonzero by assumption, and so $u + v \neq \pm(u - v)$.

(\Leftarrow) Assume that \mathcal{A} is not injective. Then there exist vectors $x, y \in \mathbb{R}^M$ such that $x \neq \pm y$ and $\mathcal{A}(x) = \mathcal{A}(y)$. Taking $S := \{n : \langle x, \varphi_n \rangle = -\langle y, \varphi_n \rangle\}$, we have $\langle x + y, \varphi_n \rangle = 0$ for every $n \in S$. Otherwise when $n \in S^c$, we have $\langle x, \varphi_n \rangle = \langle y, \varphi_n \rangle$ and so $\langle x - y, \varphi_n \rangle = 0$. Furthermore, both $x + y$ and $x - y$ are nontrivial since $x \neq \pm y$, and so neither $\{\varphi_n\}_{n \in S}$ nor $\{\varphi_n\}_{n \in S^c}$ spans \mathbb{R}^M .

Having identified an equivalent condition (the complement property) to injectivity in the real case, we now use this condition to identify the phase transition. First, we note that a spanning set for \mathbb{R}^M must have size at least M . As such, we know $\text{CP}[\Phi; \mathbb{R}^{M \times N}]$ does not hold whenever $N < 2M - 1$, since taking S to be the first $M - 1$ members will leave S^c with $\leq M - 1$ members. This suggests a phase transition of $N = 2M - 1$, but it remains to prove part (b). To get this, we first introduce the notion of full spark: An $M \times N$ matrix with $M \leq N$ is said to be *full spark* if every $M \times M$ submatrix is invertible. Note that any full spark Φ with $N \geq 2M - 1$ necessarily satisfies the complement property, since the larger of S and S^c necessarily has at least M elements, which necessarily span. As such, it suffices to show that full spark matrices form an open and dense subset. To this end, we note that the product of the determinants of all $M \times M$ submatrices forms a polynomial of the matrix entries whose zero set contains every Φ such that $\text{CP}[\Phi; \mathbb{R}^{M \times N}]$ does not hold. Moreover, this polynomial is nonzero since the Vandermonde matrix

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & N \\ \vdots & \vdots & & \vdots \\ 1^{M-1} & 2^{M-1} & \cdots & N^{M-1} \end{bmatrix}$$

is full spark (why?). The complement of this polynomial's zero set is therefore open and dense, as desired. This implies our first phase transition result:

Theorem 2 (essentially proved in [5]). $\text{Inj}[\Phi; \mathbb{R}^{M \times N}]$ exhibits a phase transition at $N = 2M - 1$.

Now that we have identified the phase transition, we consider the minimal constructions, i.e., the $M \times (2M - 1)$ real matrices which satisfy the complement property. Here, we note that for every S of size M , S^c has size $M - 1$, meaning S must index a spanning set. As such, the matrices in this extreme case are precisely the $M \times (2M - 1)$ full spark matrices. For more information about full spark matrices, see [2].

2.2 Injectivity in the Complex Case

In the previous subsection, we quickly identified a characterization of injectivity in the real case that enabled the phase transition of interest to be completely studied. The complex case appears to be a bit more involved. For example, the actual phase transition is the subject of an open conjecture, though there has been a lot of progress on this conjecture recently. Since the complex case is so much more involved, this subsection is broken into different labeled parts, concerning necessary and sufficient conditions, the phase transition, and minimal constructions.

Conditions for Injectivity in the Complex Case

We begin by defining our symbol for injectivity in the complex case:

Definition 3. $\text{Inj}[\Phi; \mathbb{C}^{M \times N}]$ denotes the statement that $\mathcal{A}: \mathbb{C}^M / \mathbb{T} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$ is injective, where $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{C}^M$.

What follows is a characterization of injectivity in the complex case:

Theorem 3 (Theorem 4 in [6]). Consider $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{C}^M$, and viewing $\{\varphi_n \varphi_n^* u\}_{n=1}^N$ as vectors in \mathbb{R}^{2M} , denote $S(u) := \text{span}_{\mathbb{R}}\{\varphi_n \varphi_n^* u\}_{n=1}^N$. Then the following are equivalent:

- (a) $\text{Inj}[\Phi; \mathbb{C}^{M \times N}]$.
- (b) $\dim S(u) \geq 2M - 1$ for every $u \in \mathbb{C}^M \setminus \{0\}$.
- (c) $S(u) = \text{span}_{\mathbb{R}}\{iu\}^\perp$ for every $u \in \mathbb{C}^M \setminus \{0\}$.

Before proving this theorem, note that unlike the characterization in the real case, it is not clear whether this characterization can be tested in finite time; instead of being a statement about all (finitely many) partitions of $\{1, \dots, N\}$, this is a statement about all $u \in \mathbb{C}^M \setminus \{0\}$. However, we can view this characterization as an analog to the real case in some sense: In the real case, the complement property is equivalent to having $\text{span}\{\varphi_n \varphi_n^* u\}_{n=1}^N = \mathbb{R}^M$

for all $u \in \mathbb{R}^M \setminus \{0\}$. As the following proof makes precise, the fact that $\{\varphi_n \varphi_n^* u\}_{n=1}^N$ fails to span all of \mathbb{R}^{2M} is rooted in the fact that more information is lost with phase in the complex case. There is actually a nice differential geometric interpretation of this result, and we will discuss it after the proof:

Proof (Proof of Theorem 3). (a) \Rightarrow (c): Suppose \mathcal{A} is injective. We need to show that $\{\varphi_n \varphi_n^* u\}_{n=1}^N$ spans the set of vectors orthogonal to iu . Here, orthogonality is with respect to the real inner product, which can be expressed as $\langle a, b \rangle_{\mathbb{R}} = \operatorname{Re} \langle a, b \rangle$. Note that

$$|\langle u \pm v, \varphi_n \rangle|^2 = |\langle u, \varphi_n \rangle|^2 \pm 2 \operatorname{Re} \langle u, \varphi_n \rangle \langle \varphi_n, v \rangle + |\langle v, \varphi_n \rangle|^2,$$

and so subtraction gives

$$|\langle u + v, \varphi_n \rangle|^2 - |\langle u - v, \varphi_n \rangle|^2 = 4 \operatorname{Re} \langle u, \varphi_n \rangle \langle \varphi_n, v \rangle = 4 \langle \varphi_n \varphi_n^* u, v \rangle_{\mathbb{R}}. \quad (1)$$

In particular, if the right-hand side of (1) is zero, then injectivity implies that there exists some ω of unit modulus such that $u + v = \omega(u - v)$. Since $u \neq 0$, we know $\omega \neq -1$, and so rearranging gives

$$v = -\frac{1 - \omega}{1 + \omega} u = -\frac{(1 - \omega)(1 + \overline{\omega})}{|1 + \omega|^2} u = -\frac{2 \operatorname{Im} \omega}{|1 + \omega|^2} iu.$$

This means $S(u)^\perp \subseteq \operatorname{span}_{\mathbb{R}}\{iu\}$. To prove $\operatorname{span}_{\mathbb{R}}\{iu\} \subseteq S(u)^\perp$, take $v = \alpha iu$ for some $\alpha \in \mathbb{R}$ and define $\omega := \frac{1 + \alpha i}{1 - \alpha i}$, which necessarily has unit modulus. Then

$$u + v = u + \alpha iu = (1 + \alpha i)u = \frac{1 + \alpha i}{1 - \alpha i}(u - \alpha iu) = \omega(u - v).$$

Thus, the left-hand side of (1) is zero, meaning $v \in S(u)^\perp$.

(b) \Leftrightarrow (c): First, (b) immediately follows from (c). For the other direction, note that iu is necessarily orthogonal to every $\varphi_n \varphi_n^* u$:

$$\langle \varphi_n \varphi_n^* u, iu \rangle_{\mathbb{R}} = \operatorname{Re} \langle \varphi_n \varphi_n^* u, iu \rangle = \operatorname{Re} \langle u, \varphi_n \rangle \langle \varphi_n, iu \rangle = -\operatorname{Re} i |\langle u, \varphi_n \rangle|^2 = 0.$$

Thus, $\operatorname{span}_{\mathbb{R}}\{iu\} \subseteq S(u)^\perp$, and by (b), $\dim S(u)^\perp \leq 1$, both of which gives (c).

(c) \Rightarrow (a): This portion of the proof is inspired by Mukherjee's analysis in [45]. Suppose $\mathcal{A}(x) = \mathcal{A}(y)$. If $x = y$, we are done. Otherwise, $x - y \neq 0$, and so we may apply (c) to $u = x - y$. First, note that

$$\langle \varphi_n \varphi_n^*(x - y), x + y \rangle_{\mathbb{R}} = \operatorname{Re} \langle \varphi_n \varphi_n^*(x - y), x + y \rangle = \operatorname{Re} (x + y)^* \varphi_n \varphi_n^*(x - y),$$

and so expanding gives

$$\begin{aligned} \langle \varphi_n \varphi_n^*(x - y), x + y \rangle_{\mathbb{R}} &= \operatorname{Re} \left(|\varphi_n^* x|^2 - x^* \varphi_n \varphi_n^* y + y^* \varphi_n \varphi_n^* x - |\varphi_n^* y|^2 \right) \\ &= \operatorname{Re} \left(-x^* \varphi_n \varphi_n^* y + \overline{x^* \varphi_n \varphi_n^* y} \right) = 0. \end{aligned}$$

Since $x + y \in S(x - y)^\perp = \operatorname{span}_{\mathbb{R}}\{i(x - y)\}$, there exists $\alpha \in \mathbb{R}$ such that $x + y = \alpha i(x - y)$, and so rearranging gives $y = \frac{1 - \alpha i}{1 + \alpha i} x$, meaning $y \equiv x \pmod{\mathbb{T}}$.

To better understand the above result, we will first develop a deeper understanding of the set \mathbb{C}^M/\mathbb{T} . If we remove zero, this set happens to be something called a *smooth manifold*, which means we can cover the set with overlapping patches, each with smooth coordinates, and with smooth coordinate transformations between the overlapping portions. To see this, consider the patches defined by

$$U_m := \{Z \in (\mathbb{C}^M \setminus \{0\})/\mathbb{T} : Z_m \neq 0\}, \quad m = 1, \dots, M.$$

We define the following coordinates over the patch U_m :

$$(z_1, \dots, z_M) = \frac{|Z_m|}{Z_m}(Z_1, \dots, Z_M),$$

where $(Z_1, \dots, Z_M) \in \mathbb{C}^M \setminus \{0\}$ denotes any representative of the corresponding point in $(\mathbb{C}^M \setminus \{0\})/\mathbb{T}$. As such, each patch has its own homeomorphism to a set of coordinates, which is an open subset of \mathbb{R}^{2M-1} (we lost a degree of freedom since the m th complex coordinate has no imaginary part). If we denote the m th homeomorphism by $f_m: U_m \rightarrow \mathbb{R}^{2M-1}$, then it is not difficult to show that each f_m , as well as the transition maps $\tau_{m,m'}: f_m(U_m \cap U_{m'}) \rightarrow f_{m'}(U_m \cap U_{m'})$ defined by $\tau_{m,m'} := f_{m'} \circ f_m^{-1}$, are all smooth.

So we now understand that $(\mathbb{C}^M \setminus \{0\})/\mathbb{T}$ is a smooth manifold with $2M-1$ real dimensions. If we consider the function \mathcal{A} over $(\mathbb{C}^M \setminus \{0\})/\mathbb{T}$, we can take its derivative at a given point u in terms of some chosen local coordinates. This amounts to taking the Jacobian at $u \in U_m$, whose rows are $\{2\varphi_n \varphi_n^* u\}_{n=1}^N$ as vectors in \mathbb{R}^{2M} , but with the column corresponding to the m th imaginary component removed. With this in mind, and using ideas from the proof of Lemma 22 in [6], Theorem 3 can be reinterpreted as follows:

Theorem 4. *\mathcal{A} is injective if and only if the derivative of \mathcal{A} is injective at every point in $(\mathbb{C}^M \setminus \{0\})/\mathbb{T}$.*

This says quite a bit about the intensity measurement mapping \mathcal{A} . Indeed, one can imagine a smooth mapping of a circle to a figure-eight curve, in which the derivative is injective at every point of the circle, but the mapping is certainly not injective. One could identify injectivity of the derivative as a local form of injectivity, and so it is rather surprising to have this be equivalent to the traditional (global) form of injectivity. Unfortunately, it is not clear what one can glean from such a feature of \mathcal{A} . Indeed, the above theorems leave a lot to be desired; compared to the complement property in the real case, it is still unclear what it takes for a complex ensemble to yield injective intensity measurements. While in pursuit of a more clear understanding, the following bizarre characterization was stumbled upon: A complex ensemble yields injective intensity measurements precisely when it yields injective *phase-only* measurements (in some sense). This is made more precise in the following theorem statement:

Theorem 5 (Theorem 5 in [6]). Consider $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{C}^M$ and the mapping $\mathcal{A}: \mathbb{C}^M/\mathbb{T} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$. Then \mathcal{A} is injective if and only if the following statement holds: If for every $n = 1, \dots, N$, either $\arg(\langle x, \varphi_n \rangle^2) = \arg(\langle y, \varphi_n \rangle^2)$ or one of the sides is not well-defined, then $x = 0$, $y = 0$, or $y \equiv x \pmod{\mathbb{R} \setminus \{0\}}$.

Proof. By Theorem 3, \mathcal{A} is injective if and only if

$$\forall x \in \mathbb{C}^M \setminus \{0\}, \quad \text{span}_{\mathbb{R}}\{\varphi_n \varphi_n^* x\}_{n=1}^N = \text{span}_{\mathbb{R}}\{ix\}^{\perp}. \quad (2)$$

Taking orthogonal complements of both sides, note that regardless of $x \in \mathbb{C}^M \setminus \{0\}$, we know $\text{span}_{\mathbb{R}}\{ix\}$ is necessarily a subset of $(\text{span}_{\mathbb{R}}\{\varphi_n \varphi_n^* x\}_{n=1}^N)^{\perp}$, and so (2) is equivalent to

$$\begin{aligned} \forall x \in \mathbb{C}^M \setminus \{0\}, \quad \text{Re}\langle \varphi_n \varphi_n^* x, iy \rangle &= 0 \quad \forall n = 1, \dots, N \\ \implies y &= 0 \text{ or } y \equiv x \pmod{\mathbb{R} \setminus \{0\}}. \end{aligned}$$

Thus, we need to determine when $\text{Im}\langle x, \varphi_n \rangle \overline{\langle y, \varphi_n \rangle} = \text{Re}\langle \varphi_n \varphi_n^* x, iy \rangle = 0$. We claim that this is true if and only if $\arg(\langle x, \varphi_n \rangle^2) = \arg(\langle y, \varphi_n \rangle^2)$ or one of the sides is not well-defined. To see this, we substitute $a := \langle x, \varphi_n \rangle$ and $b := \langle y, \varphi_n \rangle$. Then to complete the proof, it suffices to show that $\text{Im} a \bar{b} = 0$ if and only if $\arg(a^2) = \arg(b^2)$, $a = 0$, or $b = 0$.

(\Leftarrow) If either a or b is zero, the result is immediate. Otherwise, if $2\arg(a) = \arg(a^2) = \arg(b^2) = 2\arg(b)$, then 2π divides $2(\arg(a) - \arg(b))$, and so $\arg(a\bar{b}) = \arg(a) - \arg(b)$ is a multiple of π . This implies that $a\bar{b} \in \mathbb{R}$, and so $\text{Im} a\bar{b} = 0$.

(\Rightarrow) Suppose $\text{Im} a\bar{b} = 0$. Taking the polar decompositions $a = re^{i\theta}$ and $b = se^{i\phi}$, we equivalently have that $rs \sin(\theta - \phi) = 0$. Certainly, this can occur whenever r or s is zero, i.e., $a = 0$ or $b = 0$. Otherwise, a difference formula then gives $\sin \theta \cos \phi = \cos \theta \sin \phi$. From this, we know that if θ is an integer multiple of $\pi/2$, then ϕ is as well, and vice versa, in which case $\arg(a^2) = 2\arg(a) = \pi = 2\arg(b) = \arg(b^2)$. Else, we can divide both sides by $\cos \theta \cos \phi$ to obtain $\tan \theta = \tan \phi$, from which it is evident that $\theta \equiv \phi \pmod{\pi}$, and so $\arg(a^2) = 2\arg(a) = 2\arg(b) = \arg(b^2)$.

The notion of injective phase-only measurements appears similar to the notion of parallel rigidity in certain location estimation problems (for example, see [46] and references therein). It would be interesting to further investigate this relationship, but at the very least, it is rather striking that injectivity in one setting is equivalent to injectivity in the other. In [6], this equivalence is used to prove that the complement property is necessary for injectivity in the complex case. Contrary to what is claimed in [5], the first part of the proof of Theorem 1 does not suffice: It demonstrates that $u + v \neq \pm(u - v)$, but fails to establish that $u + v \not\equiv u - v \pmod{\mathbb{T}}$; for instance, it could very well be the case that $u + v = i(u - v)$, and so injectivity would not be violated in the complex case. Overall, the complement property is necessary but not sufficient

for injectivity. To see that it is not sufficient, consider measurement vectors $(1, 0)$, $(0, 1)$ and $(1, 1)$. These certainly satisfy the complement property, but $\mathcal{A}((1, i)) = (1, 1, 2) = \mathcal{A}((1, -i))$, despite the fact that $(1, i) \not\equiv (1, -i) \pmod{\mathbb{T}}$; in general, real measurement vectors fail to yield injective intensity measurements in the complex setting since they do not distinguish complex conjugates.

The theorem that follows provides one last characterization of injectivity in the complex case, and it will play a key role in our understanding of the phase transition. Before stating the result, define the real M^2 -dimensional space $\mathbb{H}^{M \times M}$ of self-adjoint $M \times M$ matrices; note that this is not a vector space over complex scalars since the diagonal of a self-adjoint matrix must be real. Given an ensemble of measurement vectors $\{\varphi_n\}_{n=1}^N \subseteq \mathbb{C}^M$, define the *super analysis operator* $\mathbf{A}: \mathbb{H}^{M \times M} \rightarrow \mathbb{R}^N$ by $(\mathbf{A}H)(n) = \langle H, \varphi_n \varphi_n^* \rangle_{\text{HS}}$; here, $\langle \cdot, \cdot \rangle_{\text{HS}}$ denotes the Hilbert-Schmidt inner product, which induces the Frobenius matrix norm. Note that \mathbf{A} is a linear operator, and yet

$$\begin{aligned} (\mathbf{A}xx^*)(n) &= \langle xx^*, \varphi_n \varphi_n^* \rangle_{\text{HS}} = \text{Tr}[\varphi_n \varphi_n^* xx^*] \\ &= \text{Tr}[\varphi_n^* xx^* \varphi_n] = \varphi_n^* xx^* \varphi_n = |\langle x, \varphi_n \rangle|^2 = (\mathcal{A}(x))(n). \end{aligned}$$

In words, the class of vectors identified with x modulo \mathbb{T} can be “lifted” to xx^* , thereby linearizing the intensity measurement process at the price of squaring the dimension of the vector space of interest; this identification has been exploited by some of the most noteworthy strides in modern phase retrieval [4, 16]. As the following lemma shows, this identification can also be used to characterize injectivity:

Theorem 6 (Lemma 9 in [6], cf. Corollary 1 in [35]). *\mathcal{A} is not injective if and only if there exists a matrix of rank 1 or 2 in the null space of \mathbf{A} .*

Proof. (\Rightarrow) If \mathcal{A} is not injective, then there exist $x, y \in \mathbb{C}^M / \mathbb{T}$ with $x \not\equiv y \pmod{\mathbb{T}}$ such that $\mathcal{A}(x) = \mathcal{A}(y)$. That is, $\mathbf{A}xx^* = \mathbf{A}yy^*$, and so $xx^* - yy^*$ is in the null space of \mathbf{A} .

(\Leftarrow) First, suppose there is a rank-1 matrix H in the null space of \mathbf{A} . Then there exists $x \in \mathbb{C}^M$ such that $H = xx^*$ and $(\mathcal{A}(x))(n) = (\mathbf{A}xx^*)(n) = 0 = (\mathcal{A}(0))(n)$. But $x \not\equiv 0 \pmod{\mathbb{T}}$, and so \mathcal{A} is not injective. Now suppose there is a rank-2 matrix H in the null space of \mathbf{A} . Then by the spectral theorem, there are orthonormal $u_1, u_2 \in \mathbb{C}^M$ and nonzero $\lambda_1 \geq \lambda_2$ such that $H = \lambda_1 u_1 u_1^* + \lambda_2 u_2 u_2^*$. Since H is in the null space of \mathbf{A} , the following holds for every n :

$$\begin{aligned} 0 &= \langle H, \varphi_n \varphi_n^* \rangle_{\text{HS}} \\ &= \langle \lambda_1 u_1 u_1^* + \lambda_2 u_2 u_2^*, \varphi_n \varphi_n^* \rangle_{\text{HS}} = \lambda_1 |\langle u_1, \varphi_n \rangle|^2 + \lambda_2 |\langle u_2, \varphi_n \rangle|^2. \end{aligned} \quad (3)$$

Taking $x := |\lambda_1|^{1/2} u_1$ and $y := |\lambda_2|^{1/2} u_2$, note that $y \not\equiv x \pmod{\mathbb{T}}$ since they are nonzero and orthogonal. We claim that $\mathcal{A}(x) = \mathcal{A}(y)$, which would complete the proof. If λ_1 and λ_2 have the same sign, then by (3), $|\langle x, \varphi_n \rangle|^2 + |\langle y, \varphi_n \rangle|^2 = 0$ for every n , meaning $|\langle x, \varphi_n \rangle|^2 = 0 = |\langle y, \varphi_n \rangle|^2$.

Otherwise, $\lambda_1 > 0 > \lambda_2$, and so $xx^* - yy^* = \lambda_1 u_1 u_1^* + \lambda_2 u_2 u_2^* = A$ is in the null space of \mathbf{A} , meaning $\mathcal{A}(x) = \mathbf{A}xx^* = \mathbf{A}yy^* = \mathcal{A}(y)$.

Of the three characterizations of injectivity in the complex case that we provided, this is by far the easiest to grasp, and perhaps due to its simplicity, much of our current understanding of the phase transition is based on this one. Still, comparing to our understanding in the real case helps to identify areas for improvement. For example, it is unclear how to test whether matrices of rank 1 or 2 lie in the null space of an arbitrary super analysis operator. Indeed, we have yet to find a “good” sufficient condition for injectivity in the complex case like the complement property or full spark provide in the real case.

The Phase Transition for Injectivity in the Complex Case

At this point, we wish to study the phase transition (presuming it exists) for $\text{Inj}[\Phi; \mathbb{C}^{M \times N}]$. To this end, we introduce the following subproblem (i.e., part (a) of the phase transition):

Problem 1. For any dimension M , what is the smallest number $N^*(M)$ of injective intensity measurements?

Interestingly, this problem has some history in the quantum mechanics literature. For example, [50] presents *Wright’s conjecture* that three observables suffice to uniquely determine any pure state. In phase retrieval parlance, the conjecture states that there exist unitary matrices U_1 , U_2 and U_3 such that $\Phi = [U_1 \ U_2 \ U_3]$ yields injective intensity measurements. Note that Wright’s conjecture actually implies that $N^*(M) \leq 3M - 2$; indeed, U_1 determines the norm (squared) of the signal, rendering the last column of both U_2 and U_3 unnecessary. Finkelstein [30] later proved that $N^*(M) \geq 3M - 2$; combined with Wright’s conjecture, this led many to believe that $N^*(M) = 3M - 2$ (for example, see [13]). However, both this and Wright’s conjecture were recently disproved in [35], in which Heinosaari, Mazzarella and Wolf invoked embedding theorems from differential geometry to prove that

$$N^*(M) \geq \begin{cases} 4M - 2\alpha(M - 1) - 3 & \text{for all } M \\ 4M - 2\alpha(M - 1) - 2 & \text{if } M \text{ is odd, } \alpha(M - 1) \equiv 2 \pmod{4} \\ 4M - 2\alpha(M - 1) - 1 & \text{if } M \text{ is odd, } \alpha(M - 1) \equiv 3 \pmod{4} \end{cases} \quad (4)$$

where $\alpha(M - 1) \leq \log_2(M)$ is the number of 1’s in the binary representation of $M - 1$; apparently, this result had previously appeared in [43, 44] as well. By comparison, Balan, Casazza and Edidin [5] proved that $N^*(M) \leq 4M - 2$, and so we at least have the asymptotic expression $N^*(M) = (4 + o(1))M$.

At this point, we should clarify some intuition for $N^*(M)$ by explaining the nature of these best known lower and upper bounds. First, the lower bound (4) follows from an older result that complex projective space \mathbb{CP}^n

does not smoothly embed into $\mathbb{R}^{4n-2\alpha(n)}$ (and other slight refinements which depend on n); this is due to Mayer [38], but we highly recommend James’s survey on the topic [36]. To prove (4) from this, suppose $\mathcal{A}: \mathbb{C}^M/\mathbb{T} \rightarrow \mathbb{R}^N$ were injective. Then \mathcal{E} defined by $\mathcal{E}(x) := \mathcal{A}(x)/\|x\|^2$ embeds \mathbb{CP}^{M-1} into \mathbb{R}^N , and as Heinosaari et al. show, the embedding is necessarily smooth; considering $\mathcal{A}(x)$ is made up of rather simple polynomials, the fact that \mathcal{E} is smooth should not come as a surprise. As such, the nonembedding result produces the best known lower bound. To evaluate this bound, first note that Milgram [40] constructs an embedding of \mathbb{CP}^n into $\mathbb{R}^{4n-\alpha(n)+1}$, establishing the importance of the $\alpha(n)$ term, but the constructed embedding does not correspond to an intensity measurement process. In order to relate these embedding results to our problem, consider the real case: It is known that for odd $n \geq 7$, real projective space \mathbb{RP}^n smoothly embeds into $\mathbb{R}^{2n-\alpha(n)+1}$ [48], which means the analogous lower bound for the real case would necessarily be smaller than $2(M-1) - \alpha(M-1) + 1 = 2M - \alpha(M-1) - 1 < 2M - 1$. This indicates that the $\alpha(M-1)$ term in (4) might be an artifact of the proof technique, rather than of $N^*(M)$.

We now consider our previous analysis of injectivity to help guide intuition about the possible phase transition. Theorem 6 indicates that we want the null space of \mathbf{A} to avoid nonzero matrices of rank ≤ 2 . Intuitively, this is easier when the “dimension” of this set of matrices is small. To get some idea of this dimension, let’s count real degrees of freedom. By the spectral theorem, almost every matrix in $\mathbb{H}^{M \times M}$ of rank ≤ 2 can be uniquely expressed as $\lambda_1 u_1 u_1^* + \lambda_2 u_2 u_2^*$ with $\lambda_1 \leq \lambda_2$. Here, (λ_1, λ_2) has two degrees of freedom. Next, u_1 can be any vector in \mathbb{C}^M , except its norm must be 1. Also, since u_1 is only unique up to global phase, we take its first entry to be nonnegative without loss of generality. Given the norm and phase constraints, u_1 has a total of $2M - 2$ real degrees of freedom. Finally, u_2 has the same norm and phase constraints, but it must also be orthogonal to u_1 , that is, $\text{Re}\langle u_2, u_1 \rangle = \text{Im}\langle u_2, u_1 \rangle = 0$. As such, u_2 has $2M - 4$ real degrees of freedom. All together, we can expect the set of matrices in question to have $2 + (2M - 2) + (2M - 4) = 4M - 4$ real dimensions.

If the set S of matrices of rank ≤ 2 formed a subspace of $\mathbb{H}^{M \times M}$ (it doesn’t), then we could expect the null space of \mathbf{A} to intersect that subspace nontrivially whenever $\dim \text{null}(\mathbf{A}) + (4M - 4) > \dim(\mathbb{H}^{M \times M}) = M^2$. By the rank-nullity theorem, this would indicate that injectivity requires

$$N \geq \text{rank}(\mathbf{A}) = M^2 - \dim \text{null}(\mathbf{A}) \geq 4M - 4. \quad (5)$$

Of course, this logic is not technically valid since S is not a subspace. It is, however, a special kind of set: a real projective variety. To see this, let’s first show that it is a real algebraic variety, specifically, the set of members of $\mathbb{H}^{M \times M}$ for which all 3×3 minors are zero. Of course, every member of S has this minor property. Next, we show that members of S are the only matrices with this property: If the rank of a given matrix is ≥ 3 , then it has an $M \times 3$ submatrix of linearly independent columns, and since the rank of its transpose is also

≥ 3 , this $M \times 3$ submatrix must have 3 linearly independent rows, thereby implicating a full-rank 3×3 submatrix. This variety is said to be projective because it is closed under scalar multiplication. If S were a projective variety over an algebraically closed field (it's not), then the projective dimension theorem (Theorem 7.2 of [34]) says that S intersects $\text{null}(\mathbf{A})$ nontrivially whenever the dimensions are large enough: $\dim \text{null}(\mathbf{A}) + \dim S > \dim \mathbb{H}^{M \times M}$, thereby implying that injectivity requires (5). Unfortunately, this theorem is not valid when the field is \mathbb{R} ; for example, the cone defined by $x^2 + y^2 - z^2 = 0$ in \mathbb{R}^3 is a projective variety of dimension 2, but its intersection with the 2-dimensional xy -plane is trivial, despite the fact that $2 + 2 > 3$.

In the absence of a proof, we pose the natural conjecture:

Conjecture 1 (The $4M - 4$ conjecture). $\text{Inj}[\Phi; \mathbb{C}^{M \times N}]$ exhibits a phase transition at $N = 4M - 4$.

As incremental progress toward solving the $4M - 4$ conjecture, we offer the following result:

Theorem 7 (Theorem 10 in [6]). *The $4M - 4$ conjecture is true when $M = 2$.*

In this case, injectivity is equivalent to \mathbf{A} having a trivial null space by Theorem 6, meaning \mathbf{A} must have rank $M^2 = 4 = 4M - 4$ for injectivity, implying part (a). For $N = 4M - 4$, \mathbf{A} has a square matrix representation, and so injectivity is equivalent to having $\det \mathbf{A} \neq 0$. As such, part (b) is proved by considering the real algebraic variety $V := \{\mathbf{A} : \text{Re det } \mathbf{A} = \text{Im det } \mathbf{A} = 0\}$ and showing that V^c is nonempty.

We can also prove the $M = 3$ case, but we first introduce Algorithm 2.2, namely the *HMW test* for injectivity; this is named after Heinosaari, Mazarella and Wolf, who implicitly introduce this algorithm in their paper [35].

Algorithm 1 The HMW test for injectivity when $M = 3$

Input: Measurement vectors $\{\varphi_n\}_{n=1}^N \subseteq \mathbb{C}^3$

Output: Whether \mathcal{A} is injective

Define $\mathbf{A}: \mathbb{H}^{3 \times 3} \rightarrow \mathbb{R}^N$ such that $\mathbf{A}H = \{\langle H, \varphi_n \varphi_n^* \rangle_{\text{HS}}\}_{n=1}^N$

if $\dim \text{null}(\mathbf{A}) = 0$ **then**

 "INJECTIVE"

{if \mathbf{A} is injective, then \mathcal{A} is injective}

else

 Pick $H \in \text{null}(\mathbf{A})$, $H \neq 0$

if $\dim \text{null}(\mathbf{A}) = 1$ and $\det(H) \neq 0$ **then**

 "INJECTIVE"

{if \mathbf{A} only maps nonsingular matrices to zero, then \mathcal{A} is injective}

else

 "NOT INJECTIVE" {in the remaining case, \mathbf{A} maps differences of rank-1 matrices to zero}

end if

end if

Theorem 8 (Theorem 11 in [6], cf. Proposition 6 in [35]). *When $M = 3$, the HMW test correctly determines whether \mathcal{A} is injective.*

Proof. First, if \mathbf{A} is injective, then $\mathcal{A}(x) = \mathbf{A}xx^* = \mathbf{A}yy^* = \mathcal{A}(y)$ if and only if $xx^* = yy^*$, i.e., $y \equiv x \pmod{\mathbb{T}}$. Next, suppose \mathbf{A} has a 1-dimensional null space. Then Lemma 6 gives that \mathcal{A} is injective if and only if the null space of \mathbf{A} is spanned by a matrix of full rank. Finally, if the dimension of the null space is 2 or more, then there exist linearly independent (nonzero) matrices A and B in this null space. If $\det(A) = 0$, then it must have rank 1 or 2, and so Lemma 6 gives that \mathcal{A} is not injective. Otherwise, consider the map

$$f: t \mapsto \det(A \cos t + B \sin t) \quad \forall t \in [0, \pi].$$

Since $f(0) = \det(A)$ and $f(\pi) = \det(-A) = (-1)^3 \det(A) = -\det(A)$, the intermediate value theorem gives that there exists $t_0 \in [0, \pi]$ such that $f(t_0) = 0$, i.e., the matrix $A \cos t_0 + B \sin t_0$ is singular. Moreover, this matrix is nonzero since A and B are linearly independent, and so its rank is either 1 or 2. Lemma 6 then gives that \mathcal{A} is not injective.

As an example, we may run the HMW test on the columns of the following matrix:

$$\Phi = \begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 0 & 1 & i \\ -1 & 0 & 0 & 1 & 1 & -1 & -2 & 2 \\ 0 & 1 & -1 & 1 & -1 & 2i & i & -1 \end{bmatrix}. \quad (6)$$

In this case, the null space of \mathbf{A} is 1-dimensional and spanned by a nonsingular matrix. As such, \mathcal{A} is injective. We are now ready to approach the $4M - 4$ conjecture in the $M = 3$ case:

Theorem 9 (Theorem 12 in [6]). *The $4M - 4$ conjecture is true when $M = 3$.*

This is proved using the HMW test. In this case, $4M - 4 = 8 = M^2 - 1$, meaning when $N = 4M - 4$, the null space of \mathbf{A} should typically be 1-dimensional. As such, the null space can be described algebraically in terms of a generalized cross product, and this is leveraged along with the HMW test to construct a real algebraic variety containing all \mathbf{A} 's which are not injective; the construction in (6) is then used to prove that the complement of this variety is nonempty, thereby proving part (b).

Recently, the American Institute of Mathematics hosted a workshop called “Frame Theory Intersects Geometry,” where experts from the two communities discussed various problems, including the $4M - 4$ conjecture. One outcome of this workshop was a paper by Conca, Edidin, Hering and Vinzant [20], which makes a major stride toward solving the $4M - 4$ conjecture:

Theorem 10 (Theorem 1.1 and Proposition 5.4 in [20]).

(a) *Part (a) of the $4M - 4$ conjecture is true whenever $M = 2^k + 1$.*

(b) Part (b) of the $4M - 4$ conjecture is true.

Proof (sketch). The proof of (a) uses certain integrality conditions, similar to the proofs of embedding results for complex projective space. Part (b) is proved using the following basic ideas: Consider the set of all $M \times N$ complex matrices Φ . This set has real dimension $2MN$. The goal is to show that the set of “bad” Φ ’s (those which fail to yield injectivity) has strictly smaller dimension. To do this, note from Theorem 6 that $\Phi = \{\varphi_n\}_{n=1}^N$ is bad precisely when there is an $M \times M$ matrix Q of rank ≤ 2 and Frobenius norm 1 such that $\varphi_n^* Q \varphi_n = 0$ for every $n = 1, \dots, N$. As such, we lift to the set of pairs (Φ, Q) which satisfy this relation. Counting the dimension of this lifted set, we note that the set of Q ’s has $4M - 5$ real dimensions, and for each Q and each n , there is a $(2M - 1)$ -dimensional set of φ_n ’s such that $\varphi_n^* Q \varphi_n = 0$. Thus, the total dimension of bad pairs (Φ, Q) is $4M - 5 + (2M - 1)N$. Recall that the bad set we care about is the set of Φ ’s for which there exists a Q such that (Φ, Q) is bad, and so we get our set by projection. Also, projections never increase the dimension of a set, and so the dimension of our set of bad Φ ’s is $\leq 4M - 5 + (2M - 1)N$. As such, to ensure that this dimension is less than the ambient dimension $2MN$, it suffices to have $4M - 5 + (2M - 1)N < 2MN$, or equivalently, $N \geq 4M - 4$. Thus, generic $M \times N$ Φ ’s with $N \geq 4M - 4$ are not bad.

Note that this result contains the previous cases where $M = 2, 3$, and the first remaining open case is $M = 4$. On my blog, I offer a US\$100 prize for a proof of the $4M - 4$ conjecture, and a can of Coca-Cola for a disproof [42].

Minimal Constructions with Injectivity in the Complex Case

In the absence of a “good” characterization of injectivity in the complex case, it is interesting to see explicit minimal constructions, as these shed some insight into the underlying structure of such ensembles. To this end, there are presently two general constructions, which we describe here.

The construction of Bodmann and Hammen [10] considers the case where the measurement vectors form a certain harmonic frame. Specifically, take the $(2M - 1) \times (2M - 1)$ discrete Fourier transform matrix and collect the first M rows; then the resulting columns are the M -dimensional measurement vectors. Note that if the original signal is known to be real, then $2M - 1$ measurements are injective whenever the measurement vectors are full spark, as this particular harmonic frame is (since all $M \times M$ submatrices are Vandermonde with distinct bases). Analogously, Bernhard and Hammen exploit the Fejer–Riesz spectral factorization theorem to say that these measurements completely determine signals from another interesting class. To be clear, identify a vector $(c_m)_{m=0}^{M-1}$ with the polynomial $\sum_{m=0}^{M-1} c_m z^m$; then Bernhard and Hammen uniquely determine the vector if the roots of the corresponding polynomial all lie outside the open unit disk in the complex plane. In general, they

actually say the polynomial is one of 2^M possibilities; here, the only ambiguity is whether a given root is at z_m or $1/\overline{z_m}$, that is, we can flip any root from outside to inside the disk. Note that this is precisely how much ambiguity we have in the real case after taking only M measurements, and in that case, we know it suffices to take only $M - 1$ additional measurements. Next, in addition to taking these $2M - 1$ measurements from before (viewed as equally spaced points on the complex unit circle), they also take $2M - 1$ measurements corresponding to equally spaced points on another unit circle in the complex plane, this one being the image of the real line under a specially chosen (think “sufficiently irrational”) Cayley map. However, this makes a total of $4M - 2$ measurements, whereas the goal is to find $4M - 4$ injective measurements – to fix this, they actually pick the second circle in such a way that it intersects the first at two points, and that these intersection points correspond to measurements from both circles, so we might as well throw two of them away.

The second known construction is due to Fickus, Nelson, Wang and the author [27]. Here, we apply two main ideas: (1) a signal’s intensity measurements with the Fourier transform is the Fourier transform of the signal’s autocorrelation; and (2) if a real, even function is sufficiently zero-padded, then it can be recovered (up to a global sign factor) from its autocorrelation. (Verifying (2) in small dimensions is a fun exercise.) In [27], we show how to generalize (2) to completely determine zero-padded complex functions, and then we identify these autocorrelations as the inverse Fourier transforms of intensity measurements with $4M - 2$ truncated and modulated discrete cosine functions. At this point, we identify certain redundancies in our intensity measurements – two of them are completely determined by the others, so we can remove them.

Considering these minimal constructions, it is striking that they both come from a construction of size $4M - 2$. This begs the following question: Does every injective ensemble of size $4M - 2$ contain an injective ensemble of size $4M - 4$?

3 Almost Injectivity

In both the real and complex cases, there appears to be a phase transition which governs how many intensity measurements are necessary and generically sufficient for injectivity. Interestingly, one can save a factor of 2 in this number of measurements by slightly weakening the desired notion of injectivity [5, 31]. To be explicit, we start with the following definition:

Definition 4. *The intensity measurement mapping \mathcal{A} is said to be almost injective if $\mathcal{A}^{-1}(\mathcal{A}(x)) = \{\omega x : |\omega| = 1\}$ for every x in an open, dense subset of \mathbb{F}^M . $\text{AlmInj}[\Phi, \mathbb{F}^{M \times N}]$ denotes the statement that the intensity measurement mapping \mathcal{A} associated with Φ is almost injective.*

This section studies the phase transition for almost injectivity. Much like injectivity, we have a much better understanding of the real case than the complex case, and we consider these separately.

3.1 Almost Injectivity in the Real Case

In this section, we start by characterizing ensembles of measurement vectors which yield almost injective intensity measurements, and similar to the characterization of injectivity, the basic idea behind the analysis is to consider sums and differences of signals with identical intensity measurements. Our characterization starts with the following lemma:

Theorem 11 (Lemma 9 in [27]). *Consider $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$ and the intensity measurement mapping $\mathcal{A}: \mathbb{R}^M / \{\pm 1\} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$. Then \mathcal{A} is almost injective if and only if almost every $x \in \mathbb{R}^M$ is not in the Minkowski sum $\text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$ for all $S \subseteq \{1, \dots, N\}$. More precisely, $\mathcal{A}^{-1}(\mathcal{A}(x)) = \{\pm x\}$ if and only if $x \notin \text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$ for any $S \subseteq \{1, \dots, N\}$.*

Proof. By the definition of the mapping \mathcal{A} , for $x, y \in \mathbb{R}^M$ we have $\mathcal{A}(x) = \mathcal{A}(y)$ if and only if $|\langle x, \varphi_n \rangle| = |\langle y, \varphi_n \rangle|$ for all $n \in \{1, \dots, N\}$. This occurs precisely when there is a subset $S \subseteq \{1, \dots, N\}$ such that $\langle x, \varphi_n \rangle = -\langle y, \varphi_n \rangle$ for every $n \in S$ and $\langle x, \varphi_n \rangle = \langle y, \varphi_n \rangle$ for every $n \in S^c$. Thus, $\mathcal{A}^{-1}(\mathcal{A}(x)) = \{\pm x\}$ if and only if for every $y \neq \pm x$ and for every $S \subseteq \{1, \dots, N\}$, either there exists an $n \in S$ such that $\langle x + y, \varphi_n \rangle \neq 0$ or an $n \in S^c$ such that $\langle x - y, \varphi_n \rangle \neq 0$. We claim that this occurs if and only if x is not in the Minkowski sum $\text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$ for all $S \subseteq \{1, \dots, N\}$, which would complete the proof. We verify the claim by seeking the contrapositive in each direction.

(\Rightarrow) Suppose $x \in \text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$. Then there exists $u \in \text{span}(\Phi_S)^\perp \setminus \{0\}$ and $v \in \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$ such that $x = u + v$. Taking $y := u - v$, we see that $x + y = 2u \in \text{span}(\Phi_S)^\perp \setminus \{0\}$ and $x - y = 2v \in \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$, which means that there is no $n \in S$ such that $\langle x + y, \varphi_n \rangle \neq 0$ nor $n \in S^c$ such that $\langle x - y, \varphi_n \rangle \neq 0$. Furthermore, u and v are nonzero, and so $y \neq \pm x$.

(\Leftarrow) Suppose $y \neq \pm x$ and for every $S \subseteq \{1, \dots, N\}$ there is no $n \in S$ such that $\langle x + y, \varphi_n \rangle \neq 0$ nor $n \in S^c$ such that $\langle x - y, \varphi_n \rangle \neq 0$. Then $x + y \in \text{span}(\Phi_S)^\perp \setminus \{0\}$ and $x - y \in \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$. Since $x = \frac{1}{2}(x + y) + \frac{1}{2}(x - y)$, we have that $x \in \text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$.

The above characterization can be simplified to form the following partial characterization of almost injectivity:

Theorem 12 (Theorem 10 in [27]). *Consider $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$ and the intensity measurement mapping $\mathcal{A}: \mathbb{R}^M / \{\pm 1\} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$. Suppose Φ spans \mathbb{R}^M and each φ_n is nonzero. Then \mathcal{A} is almost injective if and only if the Minkowski sum $\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp$ is a proper subspace of \mathbb{R}^M for each nonempty proper subset $S \subseteq \{1, \dots, N\}$.*

Note that the above result is not terribly surprising considering Theorem 11, as the new condition involves a simpler Minkowski sum in exchange

for additional (reasonable and testable) assumptions on Φ . The proof of this theorem amounts to measuring the difference between the two Minkowski sums:

Proof (Proof of Theorem 12). First note that the spanning assumption on Φ implies

$$\text{span}(\Phi_S)^\perp \cap \text{span}(\Phi_{S^c})^\perp = (\text{span}(\Phi_S) + \text{span}(\Phi_{S^c}))^\perp = \text{span}(\Phi)^\perp = \{0\},$$

and so one can prove the following identity:

$$\begin{aligned} & \text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\} \\ &= (\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \setminus (\text{span}(\Phi_S)^\perp \cup \text{span}(\Phi_{S^c})^\perp). \end{aligned} \quad (7)$$

From Theorem 11 we know that \mathcal{A} is almost injective if and only if almost every $x \in \mathbb{R}^M$ is not in the Minkowski sum $\text{span}(\Phi_S)^\perp \setminus \{0\} + \text{span}(\Phi_{S^c})^\perp \setminus \{0\}$ for any $S \subseteq \{1, \dots, N\}$. In other words, the Lebesgue measure (which we denote by $\text{Leb}[\cdot]$) of this Minkowski sum is zero for each $S \subseteq \{1, \dots, N\}$. By (7), this equivalently means that the Lebesgue measure of $(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \setminus (\text{span}(\Phi_S)^\perp \cup \text{span}(\Phi_{S^c})^\perp)$ is zero for each $S \subseteq \{1, \dots, N\}$. Since Φ spans \mathbb{R}^M , this set is empty (and therefore has Lebesgue measure zero) when $S = \emptyset$ or $S = \{1, \dots, N\}$. Also, since each φ_n is nonzero, we know that $\text{span}(\Phi_S)^\perp$ and $\text{span}(\Phi_{S^c})^\perp$ are proper subspaces of \mathbb{R}^M whenever S is a nonempty proper subset of $\{1, \dots, N\}$, and so in these cases both subspaces must have Lebesgue measure zero. As such, we have that for every nonempty proper subset $S \subseteq \{1, \dots, N\}$,

$$\begin{aligned} & \text{Leb}[(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \setminus (\text{span}(\Phi_S)^\perp \cup \text{span}(\Phi_{S^c})^\perp)] \\ & \geq \text{Leb}[\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp] - \text{Leb}[\text{span}(\Phi_S)^\perp] - \text{Leb}[\text{span}(\Phi_{S^c})^\perp] \\ & = \text{Leb}[\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp] \\ & \geq \text{Leb}[(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \setminus (\text{span}(\Phi_S)^\perp \cup \text{span}(\Phi_{S^c})^\perp)]. \end{aligned}$$

In summary, $(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \setminus (\text{span}(\Phi_S)^\perp \cup \text{span}(\Phi_{S^c})^\perp)$ having Lebesgue measure zero for each $S \subseteq \{1, \dots, N\}$ is equivalent to $\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp$ having Lebesgue measure zero for each nonempty proper subset $S \subseteq \{1, \dots, N\}$, which in turn is equivalent to the Minkowski sum $\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp$ being a proper subspace of \mathbb{R}^M for each nonempty proper subset $S \subseteq \{1, \dots, N\}$, as desired.

At this point, consider the following stronger restatement of Theorem 12: “Suppose each φ_n is nonzero. Then \mathcal{A} is almost injective if and only if Φ spans \mathbb{R}^M and the Minkowski sum $\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp$ is a proper subspace of \mathbb{R}^M for each nonempty proper subset $S \subseteq \{1, \dots, N\}$.” Note that we can move the spanning assumption into the condition because if Φ does not span, then we can decompose almost every $x \in \mathbb{R}^M$ as $x = u + v$ such that $u \in$

$\text{span}(\Phi)$ and $v \in \text{span}(\Phi)^\perp$ with $v \neq 0$, and defining $y := u - v$ then gives $\mathcal{A}(y) = \mathcal{A}(x)$ despite the fact that $y \neq \pm x$. As for the assumption that the φ_n 's are nonzero, we note that having $\varphi_n = 0$ amounts to having the n th entry of $\mathcal{A}(x)$ be zero for all x . As such, Φ yields almost injectivity precisely when the nonzero members of Φ together yield almost injectivity. With this identification, the stronger restatement of Theorem 12 above can be viewed as a complete characterization of almost injectivity. Next, we will replace the Minkowski sum condition with a rather elegant condition involving the ranks of Φ_S and Φ_{S^c} :

Theorem 13 (Theorem 11 in [27]). *Consider $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$ and the intensity measurement mapping $\mathcal{A}: \mathbb{R}^M / \{\pm 1\} \rightarrow \mathbb{R}^N$ defined by $(\mathcal{A}(x))(n) := |\langle x, \varphi_n \rangle|^2$. Suppose each φ_n is nonzero. Then \mathcal{A} is almost injective if and only if Φ spans \mathbb{R}^M and $\text{rank } \Phi_S + \text{rank } \Phi_{S^c} > M$ for each nonempty proper subset $S \subseteq \{1, \dots, N\}$.*

Proof. Considering the discussion after the proof of Theorem 12, it suffices to assume that Φ spans \mathbb{R}^M . Furthermore, considering Theorem 12, it suffices to characterize when $\dim(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) < M$. By the inclusion-exclusion principle for subspaces, we have

$$\begin{aligned} & \dim(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \\ &= \dim(\text{span}(\Phi_S)^\perp) + \dim(\text{span}(\Phi_{S^c})^\perp) - \dim(\text{span}(\Phi_S)^\perp \cap \text{span}(\Phi_{S^c})^\perp). \end{aligned}$$

Since Φ is assumed to span \mathbb{R}^M , we also have that $\text{span}(\Phi_S)^\perp \cap \text{span}(\Phi_{S^c})^\perp = \{0\}$, and so

$$\begin{aligned} & \dim(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) \\ &= \left(M - \dim(\text{span}(\Phi_S)) \right) + \left(M - \dim(\text{span}(\Phi_{S^c})) \right) - 0 \\ &= 2M - \text{rank } \Phi_S - \text{rank } \Phi_{S^c}. \end{aligned}$$

As such, $\dim(\text{span}(\Phi_S)^\perp + \text{span}(\Phi_{S^c})^\perp) < M$ precisely when $\text{rank } \Phi_S + \text{rank } \Phi_{S^c} > M$.

At this point, we point out some interesting consequences of Theorem 13. First of all, Φ cannot be almost injective if $N < M + 1$ since $\text{rank } \Phi_S + \text{rank } \Phi_{S^c} \leq |S| + |S^c| = N$. Also, in the case where $N = M + 1$, we note that Φ is almost injective precisely when Φ is full spark, that is, every size- M subcollection is a spanning set (note this implies that all of the φ_n 's are nonzero). In fact, every full spark Φ with $N \geq M + 1$ yields almost injective intensity measurements, which in turn implies that a generic Φ yields almost injectivity when $N \geq M + 1$ [5]. This is in direct analogy with injectivity in the real case; here, injectivity requires $N \geq 2M - 1$, injectivity with $N = 2M - 1$ is equivalent to being full spark, and being full spark suffices for injectivity whenever $N \geq 2M - 1$ [5]. Another thing to check is that the condition for injectivity implies the condition for almost injectivity (it does). Overall, we have the following phase transition result:

Theorem 14 (essentially proved in [5]). $\text{AlmInj}[\Phi, \mathbb{R}^{M \times N}]$ exhibits a phase transition at $N = M + 1$.

Having established that full spark ensembles of size $N \geq M + 1$ yield almost injective intensity measurements, we note that checking whether a matrix is full spark is NP-hard in general [37]. Granted, there are a few explicit constructions of full spark ensembles which can be used [2, 47], but it would be nice to have a condition which is not computationally difficult to test in general. We provide one such condition in the following theorem, but first, we briefly review the requisite frame theory.

A *frame* is an ensemble $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$ together with *frame bounds* $0 < A \leq B < \infty$ with the property that for every $x \in \mathbb{R}^M$,

$$A\|x\|^2 \leq \sum_{n=1}^N |\langle x, \varphi_n \rangle|^2 \leq B\|x\|^2.$$

When $A = B$, the frame is said to be *tight*, and such frames come with a painless reconstruction formula:

$$x = \frac{1}{A} \sum_{n=1}^N \langle x, \varphi_n \rangle \varphi_n.$$

To be clear, the theory of frames originated in the context of infinite-dimensional Hilbert spaces [22, 24], and frames have since been studied in finite-dimensional settings, primarily because this is the setting in which they are applied computationally. Of particular interest are so-called *unit norm tight frames (UNTFs)*, which are tight frames whose frame elements have unit norm: $\|\varphi_n\| = 1$ for every $n = 1, \dots, N$. Such frames are useful in applications; for example, if one encodes a signal x using frame coefficients $\langle x, \varphi_n \rangle$ and transmits these coefficients across a channel, then UNTFs are optimally robust to noise [32] and one erasure [19]. Intuitively, this optimality comes from the fact that frame elements of a UNTF are particularly well-distributed in the unit sphere [8]. Another pleasant feature of UNTFs is that it is straightforward to test whether a given frame is a UNTF: Letting $\Phi = [\varphi_1 \cdots \varphi_N]$ denote an $M \times N$ matrix whose columns are the frame elements, then Φ is a UNTF precisely when each of the following occurs simultaneously:

- (a) the rows have equal norm
- (b) the rows are orthogonal
- (c) the columns have unit norm

(This is a direct consequence of the tight frame's reconstruction formula and the fact that a UNTF has unit-norm frame elements; furthermore, since the columns have unit norm, it is not difficult to see that the rows will necessarily have norm $\sqrt{N/M}$.) In addition to being able to test that an ensemble is a UNTF, various UNTFs can be constructed using *spectral tetris* [18] (though

such frames necessarily have $N \geq 2M$), and *every* UNTF can be constructed using the recent theory of *eigensteps* [12, 28]. Now that UNTFs have been properly introduced, we relate them to almost injectivity for phase retrieval:

Theorem 15 (Theorem 12 in [27]). *If M and N are relatively prime, then every unit norm tight frame $\Phi = \{\varphi_n\}_{n=1}^N \subseteq \mathbb{R}^M$ yields almost injective intensity measurements.*

Proof. Pick a nonempty proper subset $S \subseteq \{1, \dots, N\}$. By Theorem 13, it suffices to show that $\text{rank } \Phi_S + \text{rank } \Phi_{S^c} > M$, or equivalently, $\text{rank } \Phi_S \Phi_S^* + \text{rank } \Phi_{S^c} \Phi_{S^c}^* > M$. Note that since Φ is a unit norm tight frame, we also have

$$\Phi_S \Phi_S^* + \Phi_{S^c} \Phi_{S^c}^* = \Phi \Phi^* = \frac{N}{M} I,$$

and so $\Phi_S \Phi_S^*$ and $\Phi_{S^c} \Phi_{S^c}^*$ are simultaneously diagonalizable, i.e., there exists a unitary matrix U and diagonal matrices D_1 and D_2 such that

$$UD_1 U^* + UD_2 U^* = \Phi_S \Phi_S^* + \Phi_{S^c} \Phi_{S^c}^* = \frac{N}{M} I.$$

Conjugating by U^* , this then implies that $D_1 + D_2 = \frac{N}{M} I$. Let $L_1 \subseteq \{1, \dots, M\}$ denote the diagonal locations of the nonzero entries in D_1 , and $L_2 \subseteq \{1, \dots, M\}$ similarly for D_2 . To complete the proof, we need to show that $|L_1| + |L_2| > M$ (since $|L_1| + |L_2| = \text{rank } \Phi_S \Phi_S^* + \text{rank } \Phi_{S^c} \Phi_{S^c}^*$). Note that $L_1 \cup L_2 \neq \{1, \dots, M\}$ would imply that $D_1 + D_2$ has at least one zero in its diagonal, contradicting the fact that $D_1 + D_2$ is a nonzero multiple of the identity; as such, $L_1 \cup L_2 = \{1, \dots, M\}$ and $|L_1| + |L_2| \geq M$. We claim that this inequality is strict due to the assumption that M and N are relatively prime. To see this, it suffices to show that $L_1 \cap L_2$ is nonempty. Suppose to the contrary that L_1 and L_2 are disjoint. Then since $D_1 + D_2 = \frac{N}{M} I$, every nonzero entry in D_1 must be N/M . Since S is a nonempty proper subset of $\{1, \dots, N\}$, this means that there exists $K \in (0, M)$ such that D_1 has K entries which are N/M and $M - K$ which are 0. Thus,

$$|S| = \text{Tr}[\Phi_S^* \Phi_S] = \text{Tr}[\Phi_S \Phi_S^*] = \text{Tr}[UD_1 U^*] = \text{Tr}[D_1] = K(N/M),$$

implying that $N/M = |S|/K$ with $K \neq M$ and $|S| \neq N$. Since this contradicts the assumption that N/M is in lowest form, we have the desired result.

In general, whether a UNTF Φ yields almost injective intensity measurements is determined by whether it is *orthogonally partitionable*: Φ is orthogonally partitionable if there exists a partition $S \sqcup S^c = \{1, \dots, N\}$ such that $\text{span}(\Phi_S)$ is orthogonal to $\text{span}(\Phi_{S^c})$. Specifically, a UNTF yields almost injective intensity measurements precisely when it is not orthogonally partitionable. Historically, this property of UNTFs has been pivotal to the understanding of singularities in the algebraic variety of UNTFs [25], and it has also played a key role in solutions to the Paulsen problem [9, 17]. However, it is not clear in general how to efficiently test for this property; this is why Theorem 15 is so powerful.

3.2 Almost Injectivity in the Complex Case

The complex case is not understood nearly as well as the real case, but the phase transition is arguably better understood than the one for injectivity in the complex case. However, almost injectivity hasn't received as much attention, so there are no known characterizations in the complex case, let alone "useful" ones. To begin our discussion of the phase transition, we consider the following lemma (the proof is enjoyable):

Theorem 16. *Suppose $\mathcal{A}: \mathbb{R}^P \rightarrow \mathbb{R}^N$ has a continuous Jacobian J over some open set $U \subseteq \mathbb{R}^P$. If $\text{rank}(J(x)) < P$ for every $x \in U$, then \mathcal{A} is not injective when restricted to U .*

Proof. Let z be a point in U which maximizes $\text{rank}(J(x))$, and let K denote the rank of $J(z)$. Then there are K linearly independent columns of $J(z)$ forming the submatrix $J_K(z)$. Furthermore, these columns remain linearly independent in any sufficiently small neighborhood B of z in U . (This can be established using the continuous mapping $x \mapsto \det[(J_K(x))^* J_K(x)]$.) As such, we can define the continuous mapping

$$x \mapsto P(x) := I - J_K(x)[(J_K(x))^* J_K(x)]^{-1} (J_K(x))^*$$

for all $x \in B$. By construction, $P(x)$ is the orthogonal projection onto the null space of $J(x)$. Pick some nonzero member v of the null space of $J(z)$, and consider the continuous vector field $x \mapsto P(x)v$ over B . By the Peano existence theorem, there exists $\epsilon > 0$ and $\gamma: [0, \epsilon] \rightarrow B$ such that $\gamma(0) = z$ and $\gamma'(t) = P(\gamma(t))v$ for every $t \in [0, \epsilon]$. Since $\gamma'(t) = P(\gamma(t))v$ is in the null space of $J(\gamma(t))$, we then have

$$0 = J(\gamma(t))\gamma'(t) = \frac{d}{dt}(\mathcal{A}(\gamma(t)))$$

for every $t \in [0, \epsilon]$, meaning $\mathcal{A}(x)$ is constant over all $x \in \gamma([0, \epsilon])$. Furthermore, $\gamma([0, \epsilon])$ contains more than a single point, since otherwise γ is constant, contradicting $\gamma'(0) = v \neq 0$. As such, \mathcal{A} is not injective over any sufficiently small neighborhood of z , let alone U .

Take an almost injective intensity measurement mapping \mathcal{A} and restrict it to an open set S of x 's for which $\mathcal{A}^{-1}(\mathcal{A}(x)) = \{\omega x : |\omega| = 1\}$. Note that \mathcal{A} is injective over S by assumption. Considering $(\mathbb{C}^M \setminus \{0\})/\mathbb{T}$ is a smooth manifold of real dimension $P = 2M - 1$, we can intersect S with a patch to get an open set U , and consider the Jacobian of \mathcal{A} in the patch's local coordinates. By the contrapositive of Theorem 16, we have that $N \geq \text{rank}(J(x)) \geq P = 2M - 1$ for some $x \in U$. This may lead one to believe that $\text{AlmInj}[\Phi, \mathbb{C}^{M \times N}]$ exhibits a phase transition at $N = 2M - 1$, but there is evidence to suggest that is this off by 1 due to algebraic properties of intensity measurements:

Conjecture 2. $\text{AlmInj}[\Phi, \mathbb{C}^{M \times N}]$ exhibits a phase transition at $N = 2M$.

To be clear, part (b) of this conjecture was proved by Balan, Casazza and Edidin [5], whereas a sketch of the proof of part (a) is provided in [31]. However, the latter sketch leaves much to be desired – while the argument is believable in principle, it is unclear whether their use of real algebraic geometry is sufficiently rigorous. For explicit minimal constructions in this case (assuming the conjecture is true), see [30, 31].

Acknowledgments

The author was supported by NSF Grant No. DMS-1321779. The views expressed in this chapter are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

References

1. B. Alexeev, A. S. Bandeira, M. Fickus, D. G. Mixon, Phase retrieval with polarization, *SIAM J. Imaging Sci.* 7 (2014) 35–66.
2. B. Alexeev, J. Cahill, D. G. Mixon, Full spark frames, *J. Fourier Anal. Appl.* 18 (2012) 1167–1194.
3. R. Balan, Reconstruction of signals from magnitudes of redundant representations, Available online: 1207.1134
4. R. Balan, B. G. Bodmann, P. G. Casazza, D. Edidin, Painless reconstruction from magnitudes of frame coefficients, *J. Fourier Anal. Appl.* 15 (2009) 488–501.
5. R. Balan, P. Casazza, D. Edidin, On signal reconstruction without phase, *Appl. Comput. Harmon. Anal.* 20 (2006) 345–356.
6. A. S. Bandeira, J. Cahill, D. G. Mixon, A. A. Nelson, Saving phase: Injectivity and stability for phase retrieval, to appear in *Appl. Comput. Harmon. Anal.*
7. A. S. Bandeira, Y. Chen, D. G. Mixon, Phase retrieval from power spectra of masked signals, Available online: arXiv:1303.4458
8. J. J. Benedetto, M. Fickus, Finite normalized tight frames, *Adv. Comput. Math.* 18 (2003) 357–385.
9. B. G. Bodmann, P. G. Casazza, The road to equal-norm Parseval frames, *J. Funct. Anal.* 258 (2010) 397–420.
10. B. G. Bodmann, N. Hammen, Stable phase retrieval with low-redundancy frames, Available online: arXiv:1302.5487
11. O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D. K. Satapathy, J. F. van der Veen, Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels, *Acta Cryst. A* 63 (2007) 306–314.
12. J. Cahill, M. Fickus, D. G. Mixon, M. J. Poteet, N. Strawn, Constructing finite frames of a given spectrum and set of lengths, *Appl. Comput. Harmon. Anal.* 35 (2013) 52–73.
13. E. J. Candès, Y. C. Eldar, T. Strohmer, V. Voroninski, Phase retrieval via matrix completion, *SIAM J. Imaging Sci.* 6 (2013) 199–225.

14. E. J. Candès, X. Li, Solving quadratic equations via PhaseLift when there are about as many equations as unknowns, to appear in *Found. Comput. Math.*
15. E. J. Candès, X. Li, M. Soltanolkotabi, Phase retrieval from coded diffraction patterns, Available online: [arXiv:1310.3240](https://arxiv.org/abs/1310.3240)
16. E. J. Candès, T. Strohmer, V. Voroninski, PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming, *Commun. Pure Appl. Math.* 66 (2013) 1241–1274.
17. P. G. Casazza, M. Fickus, D. G. Mixon, Auto-tuning unit norm frames, *Appl. Comput. Harmon. Anal.* 32 (2012) 1–15.
18. P. G. Casazza, M. Fickus, D. G. Mixon, Y. Wang, Z. Zhou, Constructing tight fusion frames, *Appl. Comput. Harmon. Anal.* 30 (2011) 175–187.
19. P. G. Casazza, J. Kovačević, Equal-norm tight frames with erasures, *Adv. Comput. Math.* 18 (2003) 387–430.
20. A. Conca, D. Edidin, M. Hering, C. Vinzant, An algebraic characterization of injectivity in phase retrieval, Available online: [arXiv:1312.0158](https://arxiv.org/abs/1312.0158)
21. J. C. Dainty, J. R. Fienup, Phase retrieval and image reconstruction for astronomy, In: H. Stark, ed., *Image Recovery: Theory and Application*, Academic Press, New York, 1987.
22. I. Daubechies, A. Grossmann, Y. Meyer, Painless nonorthogonal expansions, *J. Math. Phys.* 27 (1986) 1271–1283.
23. L. Demanet, P. Hand, Stable optimizationless recovery from phaseless linear measurements, to appear in *J. Fourier Anal Appl.*
24. R. J. Duffin, A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Amer. Math. Soc.* 72 (1952) 341–366.
25. K. Dykema, N. Strawn, Manifold structure of spaces of spherical tight frames, *Int. J. Pure Appl. Math.* 28 (2006) 217–256.
26. Y. C. Eldar, S. Mendelson, Phase retrieval: Stability and recovery guarantees, Available online: [arXiv:1211.0872](https://arxiv.org/abs/1211.0872)
27. M. Fickus, D. G. Mixon, A. A. Nelson, Y. Wang, Phase retrieval from very few measurements, Available online: [arXiv:1307.7176](https://arxiv.org/abs/1307.7176)
28. M. Fickus, D. G. Mixon, M. J. Poteet, N. Strawn, Constructing all self-adjoint matrices with prescribed spectrum and diagonal, *Adv. Comput. Math.* 39 (2013) 585–609.
29. J. R. Fienup, J. C. Marron, T. J. Schulz, J. H. Seldin, Hubble Space Telescope characterized by using phase-retrieval algorithms, *Appl. Optics* 32 (1993) 1747–1767.
30. J. Finkelstein, Pure-state informationally complete and “really” complete measurements, *Phys. Rev. A* 70 (2004) 052107.
31. S. T. Flammia, A. Silberfarb, C. M. Caves, Minimal informationally complete measurements for pure states, *Found. Phys.* 35 (2005) 1985–2006.
32. V. K. Goyal, M. Vetterli, N. T. Thao, Quantized overcomplete expansions in \mathbb{R}^N : Analysis, synthesis, and algorithms, *IEEE Trans. Inform. Theory* 44 (1998) 1–31.
33. R. W. Harrison, Phase problem in crystallography, *J. Opt. Soc. Am. A* 10 (1993) 1046–1055.
34. R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics, Springer, New York, 1977.
35. T. Heinosaari, L. Mazzarella, M. M. Wolf, Quantum tomography under prior information, *Commun. Math. Phys.* 318 (2013) 355–374.

36. I. M. James, Euclidean models of projective spaces, *Bull. London Math. Soc.* 3 (1971) 257–276.
37. L. Khachiyan, On the complexity of approximating extremal determinants in matrices, *J. Complexity* 11 (1995) 138–153.
38. K. H. Mayer, Elliptische Differentialoperatoren und Ganzzahligkeitssätze für charakteristische Zahlen, *Topology* 4 (1965) 295–313.
39. J. Miao, T. Ishikawa, Q. Shen, T. Earnest, Extending X-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes, *Annu. Rev. Phys. Chem.* 59 (2008) 387–410.
40. R. J. Milgram, Immersing projective spaces, *Ann. Math.* 85 (1967) 473–482.
41. R. P. Millane, Phase retrieval in crystallography and optics, *J. Opt. Soc. Am. A* 7 (1990) 394–411.
42. D. G. Mixon Saving phase: Injectivity and stability for phase retrieval, Short, Fat Matrices (blog), Available online: <http://dustingmixon.wordpress.com/2013/03/19/saving-phase-injectivity-and-stability-for-phase-r>
43. B. Z. Moroz, Reflections on quantum logic, *Int. J. Theor. Phys.* 23 (1984) 497–498.
44. B. Z. Moroz, A. M. Perelomov, On a problem posed by Pauli, *Theor. Math. Phys.* 101 (1994) 1200–1204.
45. A. Mukherjee, Embedding complex projective spaces in Euclidean space, *Bull. London Math. Soc.* 13 (1981) 323–324.
46. O. Ozyesil, A. Singer, R. Basri, Camera motion estimation by convex programming, Available online: arXiv:1312.5047
47. M. Püschel, J. Kovačević, Real, tight frames with maximal robustness to erasures, *Proc. Data Compr. Conf.* (2005) 63–72.
48. B. Steer, On the embedding of projective spaces in euclidean space, *Proc. London Math. Soc.* 21 (1970) 489–501.
49. D. L. Sun, J. O. Smith III, Estimating a signal from a magnitude spectrogram via convex optimization, Available online: arXiv:1209.2076
50. A. Vogt, Position and momentum distributions do not determine the quantum mechanical state, In: A. R. Marlow, ed., *Mathematical Foundations of Quantum Theory*, Academic Press, New York, 1978.
51. V. Voroninski, A comparison between the PhaseLift and PhaseCut algorithms, Available online: <http://math.berkeley.edu/~vladv/PhaseCutProofs.pdf>
52. I. Waldspurger, A. d’Aspremont, S. Mallat, Phase recovery, MaxCut and complex semidefinite programming, Available online: arXiv:1206.0102
53. A. Walther, The question of phase retrieval in optics, *Opt. Acta* 10 (1963) 41–49.